Influence of Holocene environmental change and anthropogenic impact on the diversity and distribution of roe deer.

Karis H. Baker & A. Rus Hoelzel

School of Biological and Biomedical Sciences, Durham University, South Road, Durham, DH1 3LE, UK

Keywords: *Capreolus capreolus*, ancient DNA, LGM, genetic diversity.

Corresponding: A. Rus Hoelzel, School of Biological and Biomedical Sciences, Durham University, South Road, Durham, DH1 3LE, UK. Fax: 0191-334-2001; email: a.r.hoelzel@dur.ac.uk.

Word count: 5,031

Abstract

Extant patterns of population structure and levels of diversity are a consequence of factors that vary in both space and time. Our objective in this study is to investigate a species that has responded to both natural and anthropogenic changes in ways that have shaped modern populations and provide insight into the key processes. The roe deer (*Capreolus capre*olus) is one of two species of deer native to Britain. During the last glacial maximum (LGM) the British habitat was largely under ice and there was a land bridge to mainland Europe. As the Earth warmed during the early Holocene the land bridge was lost. Subsequent hunting on the British mainland left the southern region extirpated of roe deer, while a refugial population remained in the north. Later reintroductions from Europe led to population expansion, especially in southern UK. Here we combine data from ancient and modern DNA to track population dynamics and patterns of connectivity, and test hypotheses about the influence of natural and anthropogenic environmental change. We find that past expansion and divergence events coincided with a warming environment and the subsequent closure of the land bridge between Europe and the UK. We also find turnover in British roe deer haplotypes between the late-Holocene and modern day, which have likely resulted from recent human disturbance activities such as habitat perturbation, over-hunting and restocking.

Introduction

Evolutionary process determines the level and structure of genetic diversity within and between natural populations. This is influenced by both spatial and temporal factors, and both adaptation and genetic drift. Understanding the relationship between specific pressures on populations and the consequences for evolutionary potential is a core objective of conservation genetics, but also an essential aspect of understanding evolution. In this study we focus on how environmental factors (including those of anthropogenic origin) have contributed to the population dynamics and distribution of a widespread terrestrial mammal, and how these processes may have helped shape the level and pattern of genetic diversity in modern populations.

Our study species is a large mammal that is currently the most common and widespread cervid in temperate habitat across Europe, the European roe deer (*Capreolus capreolus*; see Andersen et al. 1998). The European roe deer is well represented in the fossil record from the Middle to Late Pleistocene (Lister et al. 1998; Sommer and Zachos 2009), and coexisted with many other large mammal species that perished around the Pleistocene/Holocene transition (e.g. mammoth, cave bear, steppe bison, giant deer). During the Last Glacial Maximum (LGM, 23,000-18,000 YBP; Kukla et al. 2002), roe deer were likely forced into southern refugial populations along with other temperate species, later re-colonising northern Europe following climatic warming and deglaciation.

Signatures of this past history are evident in the phylogeographic patterns of modern roe deer from across Europe (for review see Sommer et al. 2009). Genetic variation of the European roe deer divides into central, eastern and western lineages. The central lineage is widespread throughout Europe, while the eastern lineage is found mainly in Greece and Serbia and the western lineage is mainly in Spain and Portugal (see Lorenzini et al. 2003; Randi et al. 2004; Lorenzini and Lovari 2006). In addition to these divisions, significant

internal structuring has been detected in roe sampled from the Italian and Iberian peninsulas. This supports the existence of a subspecies in central-southern Italy (*C. c. italicus*) (Lorenzini et al. 2002; Randi et al. 2004) and an additional Celtic–Iberian group in central-southern Spain (Royo et al. 2007). Taken together these data likely reflect the existence of several glacial refugia (Sommer et al., 2009). Following the end of the LGM roe deer re-colonised Britain via the expanse of land known as Doggerland, which once provided a direct connection to continental Europe (Yalden 1999).

Whilst climatic change has been a major force shaping the evolutionary history of the roe deer, in modern times the ever-growing impact of humans has also been significant (Baker and Hoelzel 2013). Cervids across Europe have long been under the influence of man through hunting, habitat modifications and restocking. In Britain, it was not until the last few centuries (14^{th}-18^{th}) that the impact of human activity through hunting and deforestation began to significantly affect populations of the British roe deer. These activities were believed to have caused local extinction in most areas of southern UK (Ritson 1933), and restricted roe populations to parts of Scotland and possibly some of the northern border English counties (Whitehead 1964). With the turn of the 19th century, large scale re-planting of woodland provided suitable habitat for remnant populations over much of the north to re-colonise uninhabited areas (Taylor 1948). In the south, reforestation helped facilitate the successful re-introduction of roe deer (using both native and non-native stocks) across much of southern UK (Whitehead 1964; Prior 1995; Baker and Hoelzel 2013). Reforestation and re-introductions have been so successful that roe deer populations are showing continual expansion and re-population over much of their historic range (Whitehead 1964; Ward 2005). A recent study suggested that genetic structure and diversity of British roe deer populations has been strongly influenced by recent bottlenecks and restocking activity, as well as specific aspects of life history and behaviour (Baker and Hoelzel 2013).

In this study we focus on the British population using both ancient and modern DNA, and integrate these data into published studies on roe deer genetic diversity across Europe. We test the hypothesis that historical founder and colonisation events (both natural and anthropogenic) were instrumental in defining modern patterns of diversity and population structure in this species. We also test the hypothesis that there was free movement across the land bridge at Doggerland until the land bridge was flooded, isolating two established populations either side of the British Channel. Our results provide inference about the role of historical demographic events on the modern genetic diversity of a widespread species, and reinforce earlier inference about the role of life history characteristics (Baker and Hoelzel 2013).

Materials and methods

Ancient samples (N=140) were collected from across the UK (Table 1, S1). Samples that showed amplification success are represented on a map of Britain (N = 86; Figure 1; Table S1). DNA was extracted from 0.05 g of bone powder using a QIAquick purification kit™ following the manufacturers guidelines. Precautions to avoid contamination were taken during every stage of aDNA extraction and PCR set up, which took place in a separate laboratory dedicated to ancient DNA research free from contemporary DNA or PCR product. No laboratory materials or clothing were transferred from the post amplification rooms to the ancient laboratory. All work surfaces and equipment were thoroughly cleaned with 10% bleach (sodium hypochlorite) followed by 70% ethanol. Surfaces, equipment and solutions were also routinely exposed to UV light for at least 10 minutes. All extractions and PCR work was carried out in class II PCR hoods. Negative extraction and PCR controls (1 sample in every 5) were included to detect potential contamination in reagents and cross contamination between samples. 50% of samples were replicated by extracting twice from

105 independent samples of the same bone. Additionally, all samples with unique ancient DNA
106 haplotypes were repeated with an independent PCR amplification. DNA sequences (from
107 both independent DNA extractions and PCRs) were considered authentic when independent
108 replicates from the same individual yielded identical sequences. In rare cases when
109 consistent differences between any replicates were detected, a third replicate was used to
110 verify the actual nucleotide position.
111 Two overlapping fragments of the hypervariable section of the mitochondrial control
112 region were designed to overlap with database modern sequences. The primer pairs were
113 Roe_1F: 5'-ATT ATA TGC CCC ATG CTT AT- 3' and Roe_1R:5'-CCT GAA GAA AGA
114 ACC AGA TG-3'; Roe_2F: 5'-AAC CAA GAA CTT TAC CAG- 3' and Roe_2R: 5'-GGG
115 ACA TAA TGT ACT ATG-3'. These primer pairs amplified fragments of 244 and 267 bp
116 respectively (including primers). PCR Reactions (25µl) contained 0.2 pM/µl each primer, 0.2
117 mM each dNTP, Platinum 1X High Fidelity Buffer [60 mM Tris-SO4, pH 8.9/18 mM
118 (NH4)2SO4 (Invitrogen)], 1.5 mM MgCl2, and 1 unit of Taq High Fidelity DNA polymerase
119 (Invitrogen). 2µl of DNA template was added. Amplifications were performed with the
120 following cycles: 95ºC for 5 minutes; 45 cycles at 94ºC for 45s, 51ºC (roe_1F, 1R) or 55ºC
121 (roe_2F, 2R) for 45s and 68ºC for 45 s; 68ºC for 5 min. Amplicons were sequenced in both
122 directions and PCR products were cleaned using the QIAquick PCR Purification Kit (Qiagen)
123 according to the manufacturer's instructions.
124
125 *Phylogeography, diversity and expansion signals*
126 Our ancient DNA sequence data were compared against three published modern
127 datasets (Randi et al. 2004; Royo et al. 2007); (Baker and Hoelzel 2013). Details of the exact
128 data used and corresponding accession numbers are provided in supplementary Table S4. All
129 sequences were aligned against each other using the programme CLUSTAL X (Larkin et al.

2007). A 388 bp region of the consensus mtDNA control region was used for subsequent analyses. The European dataset was divided into the lineages eastern, western and central as well as the sub-species *C. c. italicus* (as previously described by Randi et al. 2004) and the Celtic-Iberian group (as previously described by Royo et al. 2007).

Summary statistics were calculated in DnaSP10.4.9 (Rozas et al. 2003) for each of the lineages individually and combined, as well as for the ancient and modern UK samples individually and combined. The following statistics were computed: number of segregating (polymorphic) sites (S); number of unique haplotypes (h); haplotype diversity (*H*); average number of pairwise nucleotide differences (k); and nucleotide diversity (π).

The relationship between European and UK populations (using both ancient and modern data) was investigated using a median joining network (MJN) constructed using the programme NETWORK 4.5 (http://www.fluxus-engineering.com), chosen to account for possible missing nodes and alternative connections (Bandelt et al. 1999). F_{ST} and Φ_{ST} values were calculated using Arlequin v 3.0 (Excoffier et al. 2005). Two neutrality tests were performed in DnaSP: Tajima's D (Tajima 1989) and Fu's *Fs* (Fu 1997). Mismatch distributions were also used to evaluate possible events of expansion and decline (Rogers and Harpending 1992) using the sudden expansion model and goodness of fit tests (sum of squared deviations, SSD; Harpendings raggedness index R; Schneider and Excoffier 1999). Tau (τ; calculated using ARLEQUIN 3.5) was used to estimate expansion time (T) using the equation: $T = \tau/2\mu$ where μ is the mutation rate in units of substitutions per locus per generation (Rogers and Harpending 1992). The roe deer generation time was taken to be 3 years (after Randi et al. 1998; Randi et al. 2004).

153 *Demographic analyses*

154 Given that our UK ancient sample set reflected a substantial sample size, covered a
155 timeframe of over 5,000 years and was less likely to be impacted by continental
156 introductions, all demographic analyses were based on just these samples. The program
157 BEAST v 1.4.8 was used to obtain substitution parameters (based on the ancient sample set)
158 and explore past demographic change in roe deer (Drummond and Rambaut 2007).
159 Substitution rates were estimated from temporally spaced sequence data (Drummond et al.
160 2002), obtained by stratigraphic dating. For some samples, stratigraphic dates had wide date
161 ranges and therefore an average date was used across all samples. All dates were provided as
162 years before present (YBP). Input files were first generated with BEAUTi version 1.4.2.

163 Three independent MCMC runs of four chains each were run for 20,000,000 iterations,
164 of which the first 10% were discarded as burn-in. Samples from two runs (which yielded
165 similar results) were combined to estimate model parameters. Genealogies and model
166 parameters were sampled every 2,000 iterations. An explicit post mortem damage (PMD)
167 model was incorporated into each run (Rambaut et al. 2009) which takes into account the
168 potential for sequence damage to influence the outcome of the aDNA analyses. Demographic
169 inferences were essentially the same with or without the incorporation of the post-mortem
170 damage model (details available from authors). A strict molecular clock model was applied.
171 To determine the model of sequence evolution to use in this program a hierarchical likelihood
172 test in Mr. MODELTEST 2.2 (Posada and Crandall 1998) was performed. The substitution
173 model chosen was Hasegawa, Kishino and Yano (HKY) (Hasegawa et al. 1985). Independent
174 runs were combined using Tracer 1.4 (Rambaut and Drummond 2007) to generate credibility
175 intervals that represent the coalescent model and phylogenetic uncertainty and to produce
176 final estimates. For combined runs, effective sample sizes (ESSs) for each parameter

exceeded 100, which indicated efficient mixing (i.e. low autocorrelation in the Markov chain) and sufficient sampling of model parameters. The Bayesian skyline plot demographic model was applied (Drummond et al. 2005).

The program isolation with migration (IM; Hey and Nielsen 2004) was used to estimate divergence times between the ancient UK population and its most closely related European populations. To minimise substructure within sample sets, the mainland European dataset was restricted to sequences from central Europe (Germany and France), resulting in a comparison between 37 central European and 86 ancient UK sequences. The substitution rate estimated from BEAST was incorporated and the HKY model of evolution applied. Three runs were conducted using a two-step heating increment. Each Markov chain was run for 100,000,000 generations after discarding 10% burn in. The first run was conducted to determine appropriate priors for subsequent runs; unrealistic upper bounds for priors were used in this preliminary run. Uninformative priors (i.e. ranges that encompassed the entire posterior distributions) were then set for the final two runs. The final runs were conducted using identical conditions but different random number seeds to test whether multiple runs gave similar results. To ensure convergence, simulations were run until the smallest effective sample size (ESS) estimates were at least 100 (Hey 2005). Results from replicate runs did not differ, so data from the longer of the two runs are presented. The mode is reported along with the 95% HPDs (highest posterior densities; Hey and Nielsen 2004).

Direct comparison between ancient and modern populations

To compare phylogenies between modern and ancient haplotypes, median joining networks were created in NETWORK 4.5 (Bandelt et al. 1999). Inferences about possible native and non-native haplotypes were made by mapping haplotypes common to ancient and modern populations. To investigate relationships between ancient and modern populations

pairwise F_{ST} values (Weir and Cockerham 1984) were calculated for mtDNA with 1000 permutations in Arlequin 3.0 (Excoffier et al. 2005). For the calculation of F_{ST} values the ancient population could only be divided as north and south due to restrictions of small sample sizes per area, whilst the modern populations were separated by area into 6 different populations. Bonferroni correction was applied to correct for type 1 error in multiple tests.

Results

Phylogeography, genetic diversity and expansion signals

The data taken from Randi et al. (2004) and Royo et al. (2007) divided into five separate groups (thought to reflect historical geographic refugial populations); an eastern, western, central, Celtic-Iberian and *C. c. italicus* sub species. With the addition of UK samples this structure was maintained (Figure 2). The inclusion of UK data increased the number of haplotypes found across Europe from 95 (see Royo et al., 2007) to 115. Some of the haplotypes found in both ancient and modern UK populations were shared with the central European lineage, indicating a close association (Figure 2). F_{ST} and Φ_{ST} values were used to quantify the degree of differentiation among these lineages and in comparison with UK data (Table 1; Randi et al. 2004).

Overall, levels of mtDNA diversity showed a wide range within Europe (see Table 2). The central European lineage had the highest levels of diversity and the sub species *C. c. italicus* the lowest. Diversity in ancient UK populations was similar to the main European lineages, while modern UK populations were less diverse. Fu's Fs indicated a significant departure from neutrality indicating expansion for the central lineage all European lineages combined with *C. c. Italicus* and ancient UK and modern and ancient UK data combined (Table 2). The more conservative Tajima's D test indicated negative but non-significant deviations from neutrality for these same sample sets (Table 2).

The goodness of fit tests comparing an expansion model to the observed mismatch distributions revealed strong evidence for expansion for all European lineages combined with *C .c. Italicus* (SSD = 0.003, P = 0.41; R = 0.012, P = 0.47), the ancient UK population (SSD = 0.007, P = 0.68; R = 0.036, P = 0.34), modern and ancient UK combined (SSD = 0.010, P = 0.28; R = 0.035. P= 0.46) but not for the central lineage (SSD = 0.019, P = 0.00; R= 0.089, P = 0.00).

Mismatch analyses, from which τ was calculated, are represented in figure 3 a-d. Possible expansion events occurred at $\tau = 3.08$ (7,014 YBP; 95% HPD 7,919-3,932YBP) for the central lineage and at $\tau = 5.31$ (12,069 YBP; 95% HPD 19,479-4,984) for all European lineages combined with *C. c. Italicus*. For ancient UK populations the data indicate a possible expansion event at $\tau = 2.76$ (6,274 YBP; 95% HPD 10,273-2,072 YBP).

Demographic analyses

The substitution rate estimated by BSP in BEAST was 3.69×10^{-7}(95% highest posterior density interval; HPDI; 1.82×10^{-7} to 5.82×10^{-7}) substitutions per site per year (Figure S1). A coalescent reconstruction of past population dynamics (Bayesian Skyline Plot) of British roe deer based on ancient DNA shows a rapid expansion in the effective number of roe deer between 5,000 and 6,000 YBP. After that time frame, roe deer numbers appear to remain relatively stable (Figure 4).

The divergence time for ancient UK and central Europe calculated in IMa is well resolved (Table 3; Figure S2), with a posterior distribution that has a distinct peak and bounds that fall within the prior distribution. Comparing ancient UK and the central European populations, the position of the peak indicates a population split at 5,369 YBP (95% HPD 3,317-12,901 YBP; Table 3). The effective population size for ancient UK is estimated to be largest, but central Europe and ancient UK are similar (with overlapping confidence interval

252 ranges). The estimated rate of mtDNA gene flow into the UK and into Europe is very small,
253 suggesting that after separation the populations remained isolated.
254
255 *Direct comparison between ancient and modern DNA*
256 Between ancient and modern periods haplotype number, haplotype diversity, nucleotide
257 diversity and K decreased (Figure 5, Table 2). We successfully sequenced 86 (61.4%) of the
258 ancient UK samples (see supplementary S1, Table 1 for details) from which we identified 24
259 haplotypes (h/N = 0.279). From our 279 modern UK samples, 12 haplotypes were detected
260 (Figure 5, Table 2; h/N = 0.043). Six haplotypes were common to both ancient and modern
261 populations and most of these occurred in northern UK. Some haplotypes common to both
262 time points were sampled at very low frequencies and found in isolated populations (e.g. m7,
263 m8 and m11). Six haplotypes, unique only to modern populations, may represent haplotypes
264 that have either gone undetected in historical populations or that have been introduced from
265 non-native locations. One population, which likely reflects the latter scenario, is that of
266 Norfolk, which is known to be a site of non-native re-introduction, and distinctly exhibits the
267 single unique haplotype m3 (Figure 5b, Table 3). Genetic differentiation between all ancient
268 and UK contemporary populations was significant based on mitochondrial data (Table 4).
269 The most closely related contemporary population to ancient UK populations was Lancashire
270 based on mitochondrial data (the only non-significant difference). In general, mitochondrial
271 data showed that contemporary populations found in the north were more closely related to
272 ancient UK populations than those found in the south. Specifically, the southern population
273 of Norfolk was most distantly related to ancient populations.
274
275
276

Discussion

Phylogeography, genetic diversity and expansion signals

Based on a 750 bp region of the mitochondrial DNA Randi et al. (2004) identified the existence of three main roe deer lineages in Europe; central, eastern and western together with the sub species *C. c. italicus,* which each likely represent independent glacial refugia. Royo et al., (2007) later described the existence of a Celtic-Iberian group. When the UK data was combined with these European data sets, and a reduced consensus region (388bp) examined, the same structure was maintained (Figure 2). The inclusion of the UK ancient and modern genetic data revealed that 22 haplotypes were unique to Britain. The finding that all UK haplotypes cluster with the central lineage strongly supports its origins from this lineage (see Table 1 and Figure 2). Colonisation from northern Europe near the land bridge, where the central lineage is common, would be expected based on geographic proximity.

The location of the refuge from which the central lineage originated is still unknown, although a Carpathian or further eastern origin is supported by recent molecular and fossil data (see Randi et al. 2004; Lorenzini and Lovari 2006; Sommer and Zachos 2009). A colonisation route from this area would correlate well with patterns of broad-leaf forest expansion (Petit et al. 2003) which is the preferred habitat of the roe deer (Putman and Langbein 2003). Our comparisons shown in Table 1 support the affinity between the UK samples and the central lineage illustrated in the network (Figure 2), and this seems strongest for the ancient sample set. A stronger association between the ancient UK and modern European sample sets may reflect the greater inclusion of samples from translocated and bottlenecked populations among the modern UK samples. Although studies based exclusively on mtDNA data run the risk of misinterpretation due to looking only at matriline history or only a single gene tree, roe deer comparative analyses tend to show good agreement between nuclear and mtDNA indications of population structure, as seen both on a

302 fine geographic scale in the UK (Baker and Hoelzel 2013) and on a broader European scale
303 (Randi et al. 2004; Lorenzini and Lovari 2006; Sommer and Zachos 2009). As of yet there
304 are no nuclear DNA based studies which compare roe deer populations in the UK with
305 European populations.

306 Both neutrality tests and mismatch distributions suggested strong expansion events for
307 all European lineages combined, the ancient UK sample alone, and the ancient and modern
308 UK samples combined. Calculating tau (τ) from the mismatch distributions and using the
309 substitution rate calculated in BEAST, there was evidence of expansion at, 13,500 YBP for
310 the combined European lineage, and 6,300 YBP for the UK population. The expansion date
311 for the UK population was consistent with the BSP graph, which showed that UK roe
312 populations expanded over a similar timescale (Figure 4).

313 Mismatch analyses detected possible expansion events (or selective sweeps) for roe
314 deer following the LGM (see Figure 3). The LGM (23,000–18,000 YBP; Kukla et al. 2002),
315 confined roe and other temperate species to separate southern glacial refugia. This was due
316 to the permafrost and Arctic tundra ecosystems which were widespread in central Europe
317 down to a latitude of 45° (Andersen and Borns 1997). Following the LGM, species were able
318 to recolonise by expansion into formerly glaciated regions. According to the fossil record it
319 was not until the period of warming, between 14,700 – 11,600 YBP, that roe deer were able
320 to rapidly expand into much of Europe (Sommer et al. 2009; Sommer and Zachos 2009).
321 This is consistent with the estimated expansion date for the European lineages (12,000 YBP;
322 Figure 3).

323 For northern European lowlands (such as the UK) fossil evidence suggests roe did not
324 re-colonise until the early Holocene (Sommer and Zachos 2009). For the UK the earliest
325 evidence for post-glacial re-colonisation originates from Thatcham in southern England
326 where bones were radiocarbon dated to 9,439 ± 100 YBP (Sommer et al. 2009). During this

time, the central European lowlands were apparently being slowly recolonised by birch and pine woods (Usinger 2004) which would have in turn improved the environmental conditions for roe deer. However, it was not until 6,000 YBP that the vegetation pattern broadly resembled that of today (Hewitt 1999) and the expansion signal for ancient UK roe populations at 6,300 YBP (Figure 3) may have been a response to the improved environmental conditions.

Taken together our results indicate that,new habitat was quickly exploited by expanding roe deer populations following the end of the last glaciation. Randi et al. (2004) previously proposed two expansion events for continental European populations (based on 750bp segments of the mtDNA control region), scaling values of τ from their mismatch analyses with a 'phylogenetic rate' of 4-6% per million years (Myr). The resulting expansion times were estimated to have coincided with the penultimate (*c.* 250 Ka), and the last (*c.* 130 Ka) inter-glacials. Fossil evidence suggest that roe deer have been present in Europe through at least 600,000 years, since the Middle Pleistocene (Lister et al. 1998). Using the same data but with the new substitution rate calculated in this study (37% per MY) it is estimated that expansions instead occurred at 13,300 (HPD; 8,400-22,900) YBP and 8,400 YBP (HPD; 5,300-17,100) respectively. Both of these expansion dates are consistent with the European expansion signals proposed above, based on the same data, but with a reduced sequence length (388bp).

The results of this study are consistent with a number of recent studies that have replaced 'phylogenetic rates' with substitution rates directly calibrated using ancient DNA (Ho et al. 2005), more relevant to the shorter timeframe of the Holocene. For example, the substitution rate from our study (3.69×10^{-7} s.s.yr^{-1}) is similar to that for the ancient brown bear (Saarma et al. 2007) and bison (Shapiro et al. 2004) (3.2×10^{-7} and 3.0×10^{-7} s.s.yr^{-1} respectively). In fact reviews show broad consistency among a wide range of studies (Ho et

al. 2005; Ho et al. 2008). It is unlikely that all ancient DNA data sets are confounded by error, as these errors would need to have been made systematically and substantially (Ho et al. 2007) , The use of the roe deer substitution rate we calculated suggested that divergence and population expansions occur over much shorter timescales than previously proposed. As for previous studies, the more recent dates (based on higher substitution rates) are consistent with expectations based on historical environmental events (see de Bruyn et al. 2011). For the roe deer in Europe, the expansion date is consistent with expected post-glacial expansions. For the UK sample, it is consistent with the separation of the UK from the European landmass.

Demographic analyses

For the UK sample, expansion signals from IM, mismatch distributions and the BSP plot are around the time of the separation of the UK from the European landmass. When the Scandinavian and British ice sheets reached their maximum extent, and the North Sea as a consequence receded to its lowest level, Britain was connected to the continent by a land bridge (Fairbanks 1989). This dry land, referred to as Doggerland, would have allowed roe deer to migrate from central Europe. Doggerland likely existed between 8,000 (Shennan et al. 2000; Sturt et al. 2013) and 5,800 YBP (van der Molen and van Dijck 2000) when its loss resulted from warming, unlocking large quantities of water from ice caps and causing sea levels to rise. A channel restricting movement likely developed early in this period. The split may have coincided with increasingly suitable habitat in the UK, resulting in expansion of a population recently isolated from the source population in Europe (Sommer and Zachos, 2009).

Direct comparison between ancient and modern DNA

Direct comparisons between ancient and modern populations in the UK revealed an overall loss in genetic variation (see Table 2; Figure 5) which could be attributed to a recent period of bottlenecking caused by over hunting and deforestation between the late 14th and 18th centuries (Whitehead, 1964), followed by the establishment of new populations based on small founder groups. Conservation biology is often concerned with preserving native populations to secure large-scale genetic diversity and preserve possible local adaptations (Nielsen et al. 1999). In this respect, it appears that populations in the most northern parts of the UK (e.g. Scotland; Perth/Moray/Glasgow) would be most important to conserve. These populations have retained high number of 'native' haplotypes (i.e. those present in historical populations; see Figure 5 and Table S3), show a close relationship with historical populations (Table 4) and exhibit high levels of both microsatellite DNA and mtDNA variability relative to other populations (see Baker and Hoelzel 2013).

This result is concordant with the historical record which suggested that medieval bottlenecking was less intense in Scotland. For example, Whitehead (1964) claimed that roe never went extinct in this region and may have retained appreciable numbers. The other northern populations, such as those in northern England, (e.g. Carlisle, Durham, North York, Lancashire) appear to be primarily native, but lower representative numbers of haplotypes (see Table S3) may reflect near extinction or extinction in these areas followed by later re-establishment through expansion (see Baker & Hoelzel 2013). This inference is also supported by the historical record (Bewick 1790). Other areas of the UK which appear to harbour populations based on native stock include southern populations (e.g. Dorset, Wiltshire, Somerset, Berks), which are believed to have gone extinct during the medieval bottleneck and recently descended from small founder groups of introduced native stock translocated from Scotland (Whitehead, 1964). The genetic record is concordant with this

scenario, as it appears that large losses of haplotypes between ancient and modern periods have occurred in the south (see Table S2 and S3) and that the three native haplotypes currently present in these southern populations (m1, m2 and m4; see Figure 4) are also common to Scotland (see Table S3).

For other areas of the UK it appears that populations have experienced some level of influence of re- stocking involving non-native individuals. The Lancashire population was characterised by 3 native haplotypes (m1, m4, m6; see Table S3) and one potentially non-native introduced haplotype (i.e. absent from ancient populations; see Figure 4). The latter non-native haplotype may reflect the relicts of an introduction event which occurred when 12 Austrian roe were introduced into this population in 1913 to 'improve the local breed' (Whitehead 1964; Prior 1995), though it is also possible that it was present in the ancient population, but not detected. There was evidence for complete lineage replacement in Norfolk, a finding that is consistent with the records of the human translocation of non-native (German) stock into this area (see Whitehead 1964). This was supported by the findings that a single, novel haplotype unique to this location (m3; Figure 3) was detected and that this population exhibited the highest levels of differentiation (based on both mitochondrial; see Table 4 and microsatellite F_{ST} values; see Baker and Hoelzel 2013) when compared to ancient and other contemporary UK populations.

Conclusion

This study has used ancient and modern DNA to provide otherwise intractable information on the evolutionary events that have shaped European roe deer during its recent post-glacial history, including direct data on the impact of a major vicariance event (the closing of the land bridge). Re-colonisation of the roe deer across Europe seems to have occurred very rapidly as environmental conditions improved following the end of the LGM,

and our data on this process agree well with inference from the fossil records. Shortly after, roe deer re-colonised the UK but populations became isolated as the land bridge was cut from mainland Europe. Our data provide a plausible time frame for when this occurred, and suggest a post-isolation expansion. More recent anthropogenic events led to a pattern of historical bottlenecks and re-introductions, and for these events the genetic data track well with the historical data. Since establishment during the Holocene, British roe deer populations have evolved considerable genetic structure (see Baker and Hoelzel 2012), reflecting processes associated with both natural environmental and anthropogenic changes. Understanding the integration of specific responses to climatic and anthropogenic change in species of conservation concern can help predict future patterns of diversity (Hadly and Barnosky 2009), which may be fundamental to long term conservation and species management planning (Leonard 2008).

Acknowledgements

We thank Dr. Umberto Albarella at the University of Sheffield, Professor Andrew Chamberlain at the University of Manchester, Dr. Naomi Sykes at the University of Nottingham, Derek Hurst at Worcestershire Archaeological Service, Erica Macey at Birmingham Archaeology, Richard Sabin at the Natural History Museum, Peta Sadler, Huw Sherlock at Archenfield Archaeology, Catherine Smith at Scottish Urban Archaeological Trust, Ian Smith at Chester Archaeology Services, Roy Stephenson at MOLAS, and finally Dr. Rob Symmons at Fishbourne Roman Palace for the provision of ancient roe deer samples. We thank all deer stalkers and managers who helped collect modern roe deer samples as well as Hugh Rose who helped co-ordinate this process. Finally, we thank the British Deer Society and the Kenneth Whitehead Trust for funding K.B in this PhD studentship.

Conflict of interest

The authors declare no conflicts of interest.

Data archiving

Sequence data have been submitted to GenBank: accession numbers JX971589-JX971615.

457

458 **References**:
459

460 Andersen, B. G. and H. W. Borns (1997). The Ice Age world: An introduction to
461 Quatenary history and Research with Emphasis on North America and
462 Europe during the Last 2.5 Million Years. Oslo, Scandinavian University
463 Press.
464 Andersen, R., P. Duncan and J. D. C. Linnell (1998). The European Roe Deer: The
465 Biology of Success. Oslo, Scandinavian University Press.
466 Baker, K. H. and A. R. Hoelzel (2013). Evolution of population genetic structure of
467 the British roe deer by natural and anthropogenic processes (*Capreolus*
468 *capreolus*). Ecol evol **3**(1): 89-102.
469 Bandelt, H. J., P. Forster and A. Rohl (1999). Median-joining networks for inferring
470 intraspecific phylogenies. Mol Biol Evol **16**(1): 37-48.
471 Bewick, T. (1790). A general history of the quadrupeds. Newcastle upon- Tyne,
472 Hodgson, Beilby and Bewick
473 de Bruyn, M., A. R. Hoelzel, G. R. Carvalho and M. Hofreiter (2011). Faunal
474 histories from Holocene ancient DNA. Trends Ecol Evol **26**(8): 405-413.
475 Drummond, A. J., G. K. Nicholls, A. G. Rodrigo and W. Solomon (2002).
476 Estimating mutation parameters, population history and genealogy
477 simultaneously from temporally spaced sequence data. Genetics **161**(3):
478 1307-1320.
479 Drummond, A. J. and A. Rambaut (2007). BEAST: Bayesian evolutionary analysis
480 by sampling trees. BMC Evol Biol **7**.

481 Drummond, A. J., A. Rambaut, B. Shapiro and O. G. Pybus (2005). Bayesian
482 coalescent inference of past population dynamics from molecular sequences.
483 Mol Biol Evol **22**(5): 1185-1192.
484 Excoffier, L., G. Laval and S. Schneider (2005). Arlequin (version 3.0): An
485 integrated software package for population genetics data analysis.
486 Evolutionary Bioinformatics **1**: 47-50.
487 Fairbanks, R. G. (1989). A 17000-year glacio-eustatic sea level record: influence of
488 glacial melting rates on the Younger Dryas event and deep ocean circulation.
489 Nature **342**(6250): 637-642.
490 Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population
491 growth, hitchhiking and background selection. Genetics **147**(2): 915-925.
492 Hadly, E. A. and A. D. Barnosky (2009). Vertebrate fossils and the future of
493 conservation biology. Paleontological Society Papers **15**: 39-59.
494 Hasegawa, M., H. Kishino and T. A. Yano (1985). Dating the human ape spliiting by
495 a molecular clock of mitochondrial-DNA. J Mol Evol **22**(2): 160-174.
496 Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. Biol J Linn
497 Soc **68**(1-2): 87-112.
498 Hey, J. (2005). On the number of New World founders: A population genetic portrait
499 of the peopling of the Americas. PLoS Biol **3**(6): 965-975.
500 Ho, S. Y. W., S. O. Kolokotronis and R. G. Allaby (2007). Elevated substitution
501 rates estimated from ancient DNA sequences. Biology Letters **3**(6): 702-705.
502 Ho, S. Y. W., M. J. Phillips, A. Cooper and A. J. Drummond (2005). Time
503 dependency of molecular rate estimates and systematic overestimation of
504 recent divergence times. Mol Biol Evol **22**(7): 1561-1568.

505 Ho, S. Y. W., U. Saarma, R. Barnett, J. Haile and B. Shapiro (2008). The effect of
506 inappropriate calibration: three case studies in molecular ecology. Plos One
507 **3**(2): e1615.
508 Kukla, G. J., A. C. Clement, M. A. Cane, J. E. Gavin and S. E. Zebiak (2002). Last
509 interglacial and early glacial ENSO. Quaternary Research **58**(1): 27-31.
510 Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H.
511 McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson,
512 T. J. Gibson and D. G. Higgins (2007). Clustal W and clustal X version 2.0.
513 Bioinformatics **23**(21): 2947-2948.
514 Leonard, J. A. (2008). Ancient DNA applications for wildlife conservation. Mol Ecol
515 **17**(19): 4186-4196.
516 Lister, A. M., P. Grubb and S. R. M. Sumner (1998). Taxonomy, morphology and
517 evolution of European roe deer. The European Roe Deer: The Biology of
518 Success. P. D. a. J. D. C. L. R.Andersen. Oslo, Norway, Scandinavian
519 University Press 23-46.
520 Lorenzini, R. and S. Lovari (2006). Genetic diversity and phylogeography of the
521 European roe deer: the refuge area theory revisited. Biol J Linn Soc **88**(1):
522 85-100.
523 Lorenzini, R., S. Lovari and M. Masseti (2002). The rediscovery of the Italian roe
524 deer: genetic differentiation and management implications. Italian Journal of
525 Zoology **69**(4): 367-379.
526 Lorenzini, R., C. San Jose, F. Braza and S. Aragon (2003). Genetic differentiation
527 and phylogeography of roe deer in Spain, as suggested by mitochondrial
528 DNA and microsatellite analysis. Italian Journal of Zoology **70**(1): 89-99.

529 Nielsen, E. E., M. M. Hansen and V. Loeschcke (1999). Analysis of DNA from old
530 scale samples: technical aspects, applications and perspectives for
531 conservation. Hereditas **130**(3): 265-276.
532 Petit, R. J., I. Aguinagalde, J. L. de Beaulieu, C. Bittkau, S. Brewer, R. Cheddadi, R.
533 Ennos, S. Fineschi, D. Grivet, M. Lascoux, A. Mohanty, G. M. Muller-
534 Starck, B. Demesure-Musch, A. Palme, J. P. Martin, S. Rendell and G. G.
535 Vendramin (2003). Glacial refugia: Hotspots but not melting pots of genetic
536 diversity. Science **300**(5625): 1563-1565.
537 Posada, D. and K. A. Crandall (1998). MODELTEST: testing the model of DNA
538 substitution. Bioinformatics **14**(9): 817-818.
539 Prior, R. (1995). The Roe Deer- Conservation of a Native Species, Swan- Hill Press.
540 Putman, R. and J. Langbein (2003). The Deer Manager's Companion; A Guide to the
541 Management of Deer in the Wild and in Parks. Shrewsbury, Swan Hill Press.
542 Rambaut, A. and A. J. Drummond (2007). Tracer v 1.4. University of Edinburgh,
543 UK.
544 Rambaut, A., S. Y. W. Ho, A. J. Drummond and B. Shapiro (2009). Accommodating
545 the effect of ancient DNA damage on inferences of demographic histories.
546 Mol Biol Evol **26**(2): 245-248.
547 Randi, E., P. C. Alves, J. Carranza, S. Milosevic-Zlatanovic, A. Sfougaris and N.
548 Mucci (2004). Phylogeography of roe deer (*Capreolus capreolus*)
549 populations: the effects of historical genetic subdivisions and recent
550 nonequilibrium dynamics. Mol Ecol **13**(10): 3071-3083.

551 Randi, E., M. Pierpaoli and A. Danilkin (1998). Mitochondrial DNA polymorphism
552 in populations of Siberian and European roe deer (*Capreolus pygargus* and
553 *Capreolus capreolus*). Heredity **80**: 429-437.
554 Ritson, G. (1933). The Roe deer in Cumberland. Natural history Society: Carlisle.
555 Rogers, A. R. and H. Harpending (1992). Population- growth makes waves in the
556 distribution of pairwise genetic- differences Mol Biol Evol **9**(3): 552-569.
557 Royo, L. J., G. Pajares, I. Alvarez, I. Fernandez and F. Goyache (2007). Genetic
558 variability and differentiation in Spanish roe deer (*Capreolus capreolus*): A
559 phylogeographic reassessment within the European framework. Mol
560 Phylogenet Evol **42**(1): 47-61.
561 Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer and R. Rozas (2003). DnaSP,
562 DNA polymorphism analyses by the coalescent and other methods.
563 Bioinformatics **19**(18): 2496-2497.
564 Saarma, U., S. Y. W. Ho, O. G. Pybus, M. Kaljuste, I. L. Tumanov, I. Kojola, A. A.
565 Vorobiev, N. I. Markov, A. P. Saveljev, H. Valdmann, E. A. Lyapunova, A.
566 V. Abramov, P. Mannil, M. Korsten, E. Vulla, S. V. Pazetnov, V. S.
567 Pazetnov, S. V. Putchkovskiy and A. M. Rokov (2007). Mitogenetic structure
568 of brown bears (*Ursus arctos*) in northeastern Europe and a new time frame
569 for the formation of European brown bear lineages. Mol Ecol **16**(2): 401-413.
570 Schneider, S. and L. Excoffier (1999). Estimation of past demographic parameters
571 from the distribution of pairwise differences when the mutation rates very
572 among sites: Application to human mitochondrial DNA. Genetics **152**(3):
573 1079-1089.

574 Shapiro, B., A. J. Drummond, A. Rambaut, M. C. Wilson, P. E. Matheus, A. V.
575 Sher, O. G. Pybus, M. T. P. Gilbert, I. Barnes, J. Binladen, E. Willerslev, A.
576 J. Hansen, G. F. Baryshnikov, J. A. Burns, S. Davydov, J. C. Driver, D. G.
577 Froese, C. R. Harington, G. Keddie, P. Kosintsev, M. L. Kunz, L. D. Martin,
578 R. O. Stephenson, J. Storer, R. Tedford, S. Zimov and A. Cooper (2004).
579 Rise and fall of the Beringian steppe bison. Science **306**(5701): 1561-1565.
580 Shennan, I., K. Lambeck, R. Flather, B. Horton, J. McCartur, J. Innes, J. Lloyd, M.
581 Rutherford and R. Wingfield (2000). Modelling western North Sea
582 palaeogeo-graphies and tidal changes during the Holocene. Holocene Land-
583 Ocean Inter-action and Environmental Change around the North Sea.
584 I.Shennan and J. Andrews. Geological Society, London, Special publications,
585 **166**, 299-319.
586 Sommer, R. S., J. M. Fahlke, U. Schmolcke, N. Benecke and F. E. Zachos (2009).
587 Quaternary history of the European roe deer *Capreolus capreolus*. Mammal
588 Rev **39**(1): 1-16.
589 Sommer, R. S. and F. E. Zachos (2009). Fossil evidence and phylogeography of
590 temperate species: 'glacial refugia' and post-glacial recolonization. J Biogeogr
591 **36**(11): 2013-2020.
592 Sturt, F., D. Garrow and S. Bradley (2013). New models of North West European
593 Holocene palaeogeography and inundation. J Archaeol Sci(40): 3963-3976.
594 Tajima, F. (1989). The effect of change in population size on DNA polymorphism
595 Genetics **123**(3): 597-601.
596 Taylor, W. L. (1948). The distribution of wild deer in England and Wales. J Anim
597 Ecol **17**: 151-164.

598 Usinger, H. (2004). Vegetation and climate of the lowlands of northern Central
599 Europe and adjacent areas around the Younger Dryas – Preboreal transition –
600 with special emphasis on the Preboreal oscillation. Hunters in a Changing
601 World. Environment and Archaeology of the Pleistocene-Holocene
602 Transition (ca. 11 000–9000 B.C.) in Northern Europe. T. T. a. B. V. Eriksen.
603 Rahden, Marie Leidorf Publisher**:** 1-26.
604 van der Molen, J. and B. van Dijck (2000). The evolution of the Dutch and Belgian
605 coasts and the role of sand supply from the North Sea. Global and Planetary
606 Change **27**(1-4): 223-244.
607 Ward, A. I. (2005). Expanding ranges of wild and feral deer in Great Britain.
608 Mammal Rev **35**(2): 165-173.
609 Weir, B. S. and C. C. Cockerham (1984). Estimating f-statistics for the analysis of
610 population- structure Evolution **38**(6): 1358-1370.
611 Whitehead, G. K. (1964). The Deer of Great Britain and Ireland. London, Routledge
612 & Kegan Paul.
613 Yalden, D. W. (1999). The History of British Mammals. London, Poyser.

614

615

Figure legends

Figure 1. Map showing the locations from where modern (black circles) and ancient (grey circles) mtDNA sequences originated from. The sample numbers are shown either alongside or within each sample circle.

Figure 2. Median joining network (MJN) computed using 115 haplotypes from continental European data sets along with the ancient and modern UK data sets. Circle size is proportional to haplotype frequencies. The individual coloured circles represent the lineages, sub-species and groups formerly defined by Randi et al. (2004) and Royo et al. (2007) where central lineage = yellow, eastern lineage = green, western lineage = pink, the sub species *C. c. italicus* = white, the Celtic-Iberian group = pink and the UK populations from this study where ancient UK = grey and modern UK = black.

Figure 3. Mismatch distributions for mitochondrial DNA haplotypes sampled from; the central lineage (a), all European lineages and *C. c .italicus* combined (b), ancient UK (c) and ancient and modern UK populations combined (d).

Figure 4. A Bayesian skyline plot derived from ancient UK roe deer mtDNA d-loop sequences. The x axis is in units of years before present, and the y axis is equal to population size (the product of the effective population size and the generation length in years = 3). The black line is the median estimate, and the blue area shows the 95% highest posterior density intervals.

639 Figure 5. A direct comparison of ancient (a) versus modern (b) median joining
640 networks computed in NETWORK. In (a) grey circles indicate ancient haplotypes
641 (n= 24; a1-24 see supplementary). In (b) black circles indicate modern haplotypes
642 (n=12; m1-12; see Table supplementary). For both networks the asterisks (*) denote
643 those haplotypes that are common to both ancient and modern populations

644 Table 1. Pairwise F_{ST}s (above diagonal) and Φ_{ST} (below diagonal) for roe deer between European
645 lineages, *C. c. Italicus*, the Celtic-Iberian group and ancient and modern UK populations for 388 bp
646 of the mt-DNA control region. Values in bold indicate significance after Bonferroni correction.

647 Table 2. Population genetic summary and demographic statistics for European and UK populations;
648 n: number of individuals; h: number of haplotypes; *H*: haplotype diversity (s.d.); π: nucleotide
649 diversity; k: average number of nucleotide differences; *Fs*: Fu's *Fs*; D: Tajima's D.

650 Table 3. Maximum likelihood estimates and 95% highest posterior density (HPD) intervals (in
651 parentheses) of isolation and migration model parameters and their respective demographic
652 conversions for the UK population and central Europe. The model parameters given in italics (*t* & *m*)
653 are scaled by μ. The demographic parameters (not italicised) are based on an estimate of μ (see text)
654 where: Ne = effective population size; t = divergence time in years; and m = average number of
655 migrants per 1000 generations per gene copy.
656
657 Table 4**Error! No text of specified style in document.** Pairwise F_{ST} values based on a portion of the
658 mt-DNA d loop between UK roe from six contemporary populations roe and ancient north and south
659 samples. Values in bold indicate significance after Bonferroni adjustment.

660 Table 1

	1	2	3	4	5	6	7
1. Central lineage	0	**0.11**	**0.09**	**0.31**	**0.19**	**0.06**	**0.13**
2. Western lineage	**0.55**	0	**0.12**	**0.39**	**0.22**	**0.13**	**0.19**
3. Eastern lineage	**0.58**	**0.68**	0	**0.34**	**0.20**	**0.11**	**0.18**
4. *C.c. Italicus*	**0.39**	**0.75**	**0.65**	0	**0.50**	**0.39**	**0.40**
5. Celtic-Iberian lineage	**0.51**	**0.73**	**0.52**	**0.83**	0	**0.22**	**0.28**
6. Ancient UK	**0.09**	**0.56**	**0.58**	**0.61**	**0.59**	0	**0.09**
7. Modern UK	**0.15**	**0.61**	**0.61**	**0.60**	**0.57**	**0.10**	0

661 Table 2.

	n	h	hd ± sd	π	k	F's	P	D	P
Central lineage	394	44	0.91 ± 0.0006	0.0075	2.98	**-26.94**	**0.000**	-1.12	>0.10
Eastern Lineage	188	22	0.89 ± 0.0001	0.0080	3.20	**-4.75**	**0.05**	-0.32	>0.10
Western lineage	90	15	0.87 ± 0.014	0.0068	2.63	**-3.07**	**0.06**	0.13	>0.10
Celtic-Iberian lineage	59	6	0.68 ± 0.035	0.0025	0.96	-0.99	0.143	-1.88	>0.10
C.c. Italicus	105	5	0.36 ± 0.0570	0.0009	0.39	-2.10	0.075	-0.28	>0.10
All European lineages and *C.c. Italicus*	836	92	0.962 ±0.002	0.0130	4.86	**-34.44**	**0.000**	-0.96	>0.10
Ancient UK	86	24	0.88 ± 0.0004	0.0069	2.76	**-12.40**	**0.000**	-1.34	>0.10
Modern UK	279	12	0.76 ± 0.0003	0.0062	2.47	0.32	0.131	0.43	>0.10
Ancient and Modern UK combined	366	30	0.82 ± 0.0130	0.0067	2.64	**-11.75**	**0.000**	-1.25	>0.10

662

663

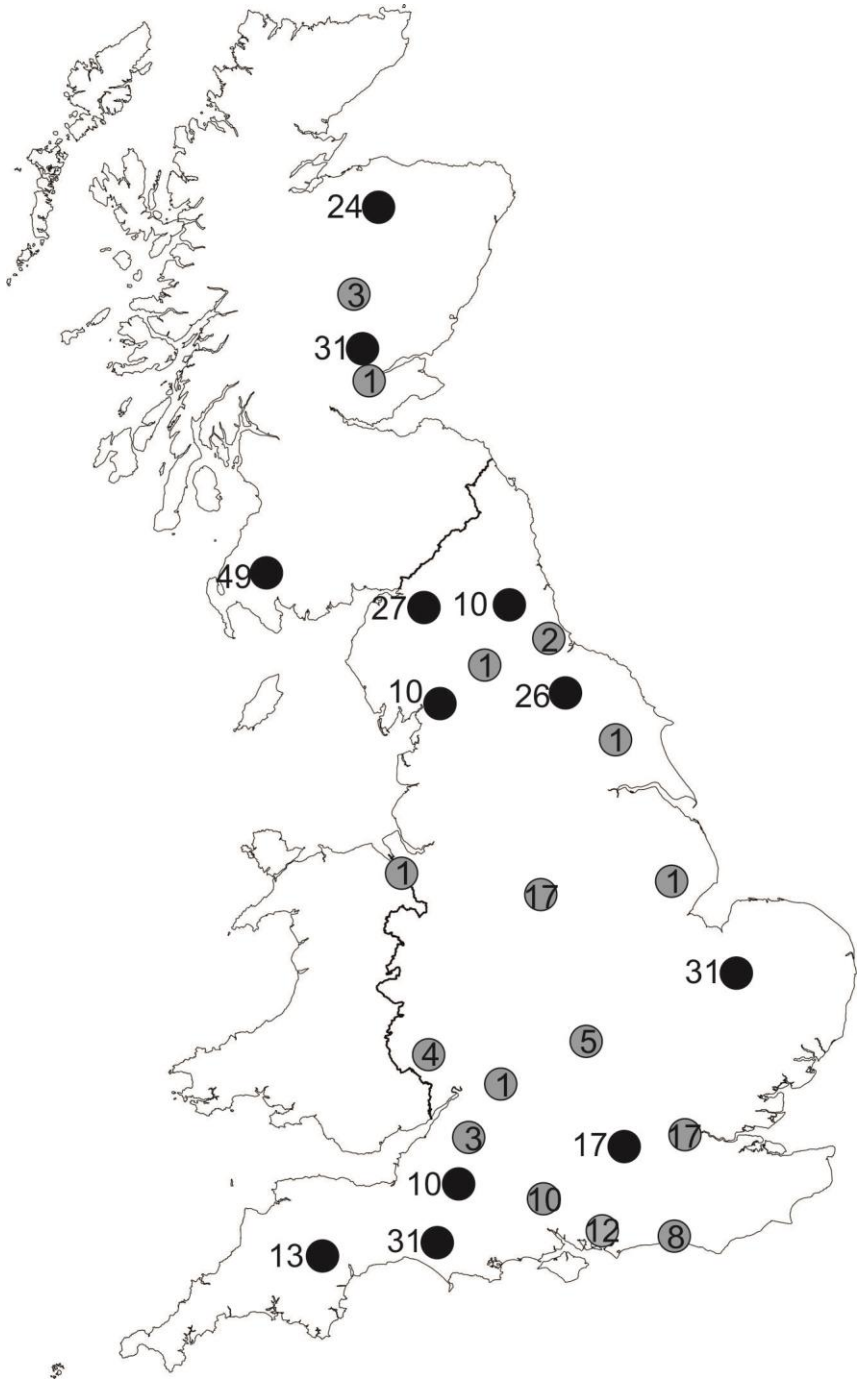
Table 3.

Parameter	**Ancient UK**	**Central Europe**
t	0.84 (0.52-2.01)	
t	5368 (3317-12868)	
Ne	32676 (18414-59616)	13216 (4595-39205)
Ne (ancestral)	2313 (728-14293)	
m	0.01 (0.01-1.07)	0.01 (0.03-4.05)
m	0.0047 (0.0047-0.500)	0.0047 (0.014-1.89)

664 Table 4.

	1	2	3	4	5	6	7
1. Norfolk							
2. Hamps, Somerset, Wilts and Berkshire	**0.88**						
3. North York, Durham and Carlisle	**0.88**	**0.58**					
4. Perth and Moray	**0.79**	**0.34**	**0.52**				
5. Ayrshire	**0.89**	**0.63**	0.007	**0.56**			
6. Lancashire	**0.86**	**0.29**	**0.41**	**0.25**	**0.44**		
7. Ancient south	**0.66**	**0.24**	**0.51**	**0.10**	**0.53**	0.08	
8. Ancient north	**0.79**	**0.26**	**0.2**	**0.17**	**0.25**	0.017	**0.167**

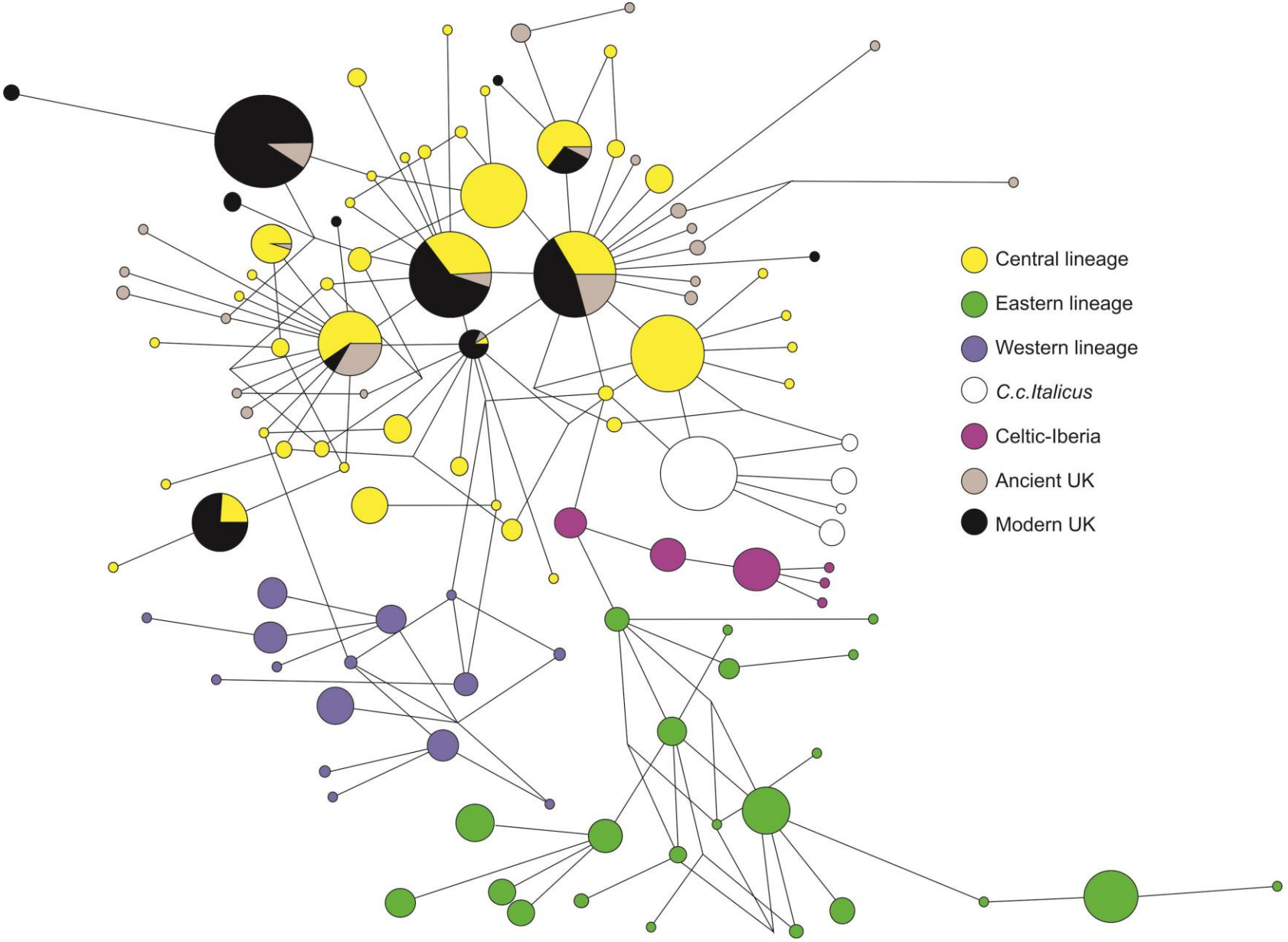
665 Figure 1.



666

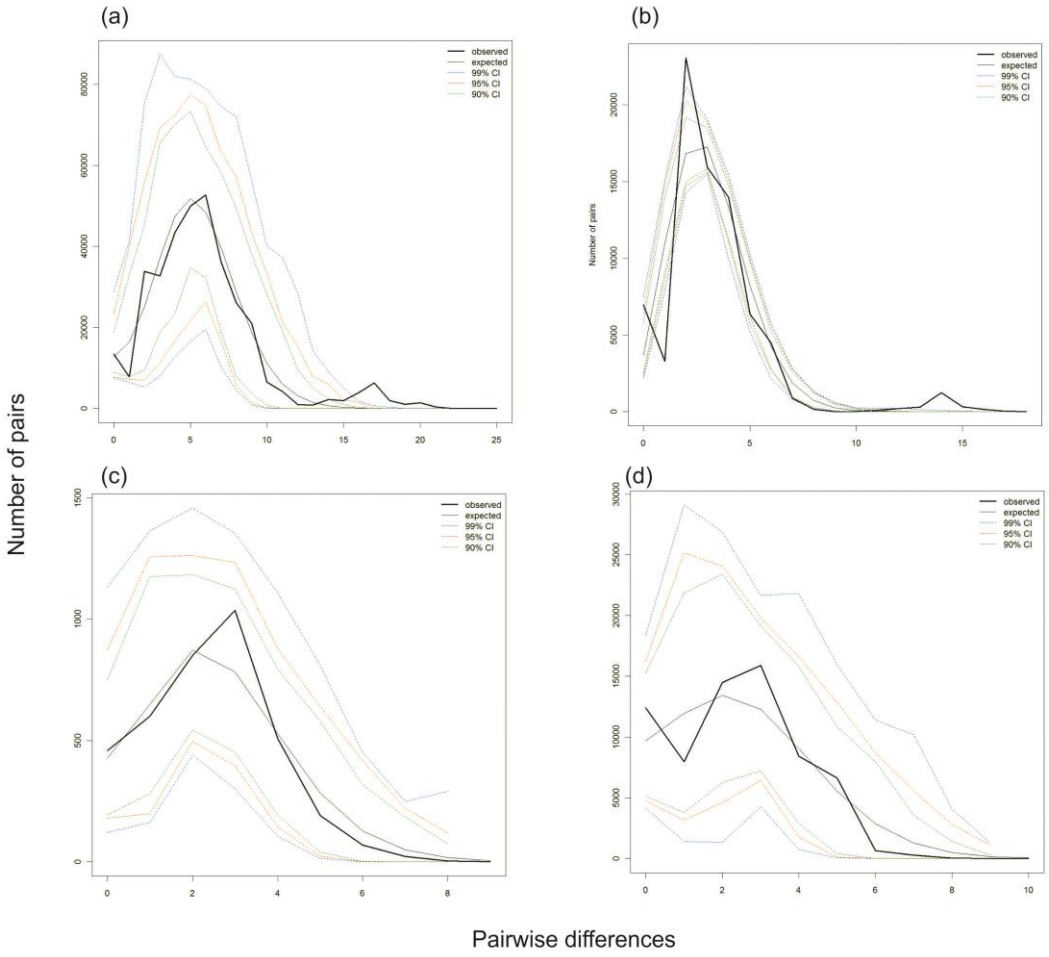
667

668

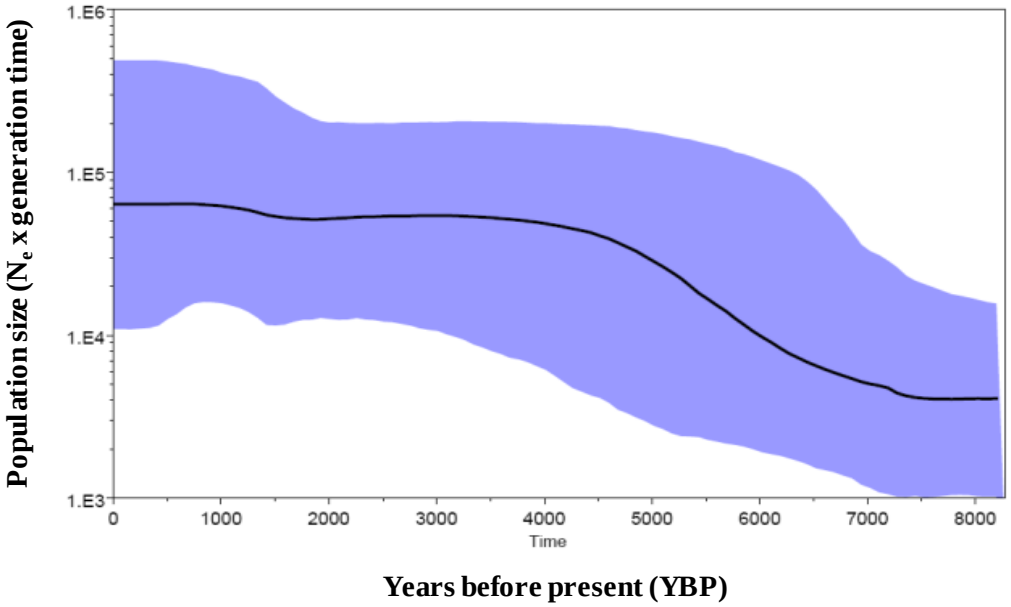

Central lineage
Eastern lineage
Western lineage
C.c.Italicus
Celtic-Iberia
Ancient UK
Modern UK

669

Figure 3.


(a)
(b)
(c)
(d)
observed
expected
99% CI
95% CI
90% CI
Number of pairs
Pairwise differences

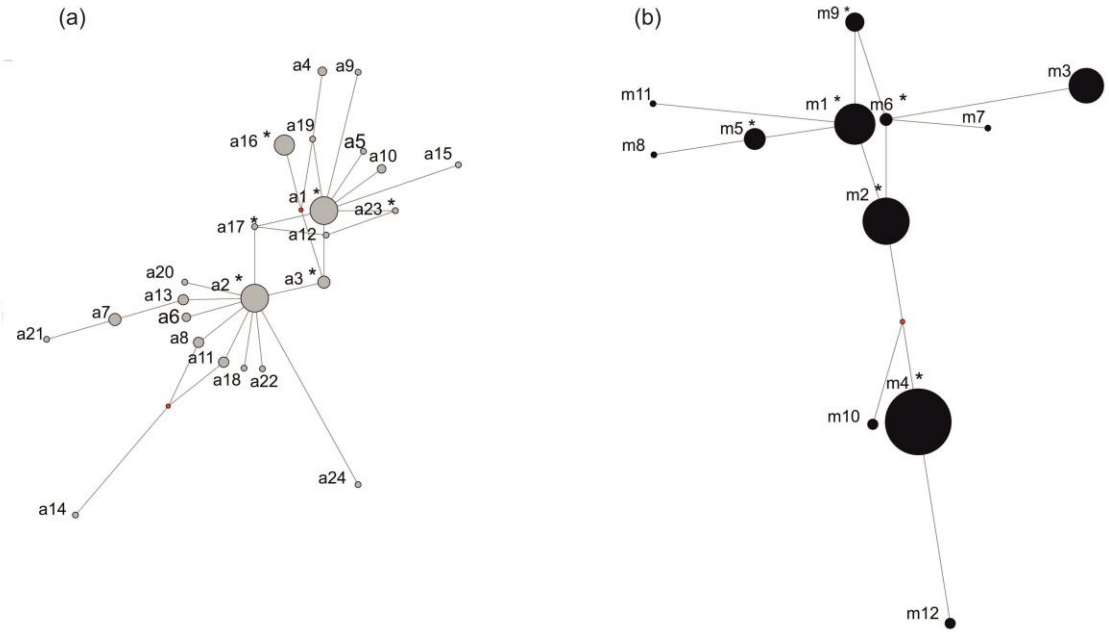
674 Figure 4.



675

676

677 Figure 5.



678

679

680

681

682

683

Supplementary

S1, Table 1. Detailed description of samples including the location, site and site code origin of each sample, the number of samples extracted (N extracted) and successfully amplified for DNA (N successful) and, finally, the approximate date range of samples (provided from stratigraphic information).

Location	Site	Site code	N extracted	N succesful	Approximate date range (AD/BC)	Period
London	Moorgate	MRG	2	2	145-170 AD	Roman
	Wood Street	WOO	1	1	190-400 AD	Roman
	Wood Street	WOO	1	1	1050-1100 AD	Norman
	Baltic Exchange	BAX	2	2	100-250 AD	Roman
	Baltic Exchange	BAX	1	1	250-400 AD	Roman
	Fenchurch Street	FEH	2	2	1050-1150 AD	Norman
	Fenchurch Street	FEH	1		50-80 AD	Roman
	London Bridge	LBI	4	4	70-160 AD	Roman
	Regis House and Ridegway House	KWS	1	1	no date	Unknown
	Upper Thames Street	UP	1	1	970-1050 AD	Late Anglo Saxon
	Borough High Street	BGH	2	2	100-160 AD	Roman
	Walthamstow	WAL	3			Pleistocene
Kent	Bishopstone	BIS	8	8	900-1000 AD	Late Anglo Saxon
Oxfordshire	Banbury Castle	BAN	7	4	1095-1292 AD	Late Anglo Saxon/early Medieval
Gloucestershire	Salmonsbury Camp	SAL	1	1	700 BC - 43 AD	Iron age
Hampshire	Faccombe Netherton	FAC	7	4	850-1070 AD	Anglo Saxon
	Faccombe Netherton	FAC	9	4	1066-1154 AD	Norman
	Faccombe Netherton	FAC	5	2	1154-1485 AD	Medieval

S1, Table 1 continued. Detailed description of samples including the location, site and site code origin of each sample, the number of samples extracted (N extracted) and successfully amplified for DNA (N successful) and, finally, the approximate date range of samples (provided from stratigraphic information).

Location	Site	Site code	N extracted	N succesful	Approximate date range (AD/BC)	Period
Wiltshire	Durrington Wells	DW	5	3	2600 BC	Late Neolithic
	Boscombe Down	BD	1		2300 - 700 BC	Bronze Age
Sussex	Fishbourne Roman Palace	FB	12	12	45-180 AD	Roman
	Whitehawk Camp	WC	4		4000 -2500 BC	Neolithic
Somerset	Glastonbury Lake Village	GLV	3		700 - 400 BC	Early Iron Age
Hereford	Cathedral House	CH	3	3	1100-1200 AD	Medieval
	Gaol Street	GAO	3	2	1066-1485 AD	Medieval
Chester	Unknown	CHE	1	1	1200-1350 AD	Medieval
Lincolnshire	Welland Bank Quarry	WBQ	4	1	1300 - 700 BC	Bronze Age
Derbyshire	Carsington cave	CPC	14	14	5678 - 3447 BC	Neolithic
	Carsington cave	CPC	2	2	1248- 630 BC	Late Bronze age/early Iron Age
Yorkshire	Staple Howe	STP	9	1	700 BC - 43 AD	Iron Age
Durham	Barnard Castle	BAR	1	1	1095-1292 AD	Medieval
	Arbeia Roman Fort	ARB	2	2	200-350 AD	Roman
Perthshire	Horse Cross	HC	3	3	1400 AD	Medieval
	Horse Cross	HC	2		1100-1250 AD	Norman
	Holyrod, Edinburgh	HLY	1	1	1500 AD	Medieval
Northumberland	Roman Vindolanda	VIN	12	0		Roman
	Total		**140**	**86**		

694 S2, Table 2. Ancient haplotypes. GenBank accession numbers for corresponding haplotypes are ******(to be input upon acceptance).

Haplotype	n individuals	List of haplotypes with location and site
a1	19	London (FEH, WOO, UP, BGH), Derbyshire (CPC), Sussex (FIS), Kent (BIS), Hampshire (FAC) and Oxfordshire (BAN)
a2	20	London (WOO, LBI, KWS, BGH), Derbyshire (CPC), Sussex (FIS), Kent (BIS), Hampshire (FAC), Hereford (GAO), Wiltshire (DW), Durham (ARB)
a3	4	London (MRG), Durham (ARB), Hampshire (FAC)
a4	2	Hereford (GAO)
a5	1	Sussex (FIS)
a6	2	Sussex (FIS)
a7	3	Sussex (FIS) and Kent (BIS)
a8	4	London (BAX, LBI) and Sussex (FIS)
a9	1	Sussex (FIS)
a10	2	Sussex (FIS) and Hampshire (FAC)
a11	3	London (FEH), Hampshire (FAC) and Oxfordshire (BAN)
a12	1	London (BAX)
a13	3	London (MRG), Perthshire (HC), Durham (BAR)
a14	1	Hereford (CH)
a15	1	Chester (CHE)
a16	11	Derbyshire (CPC) and Lincolnshire (WQ)
a17	1	Perthshire (HC)
a18	1	Wiltshire (DW)
a19	1	Yorkshire (STP)
a20	1	Hampshire (FAC)
a21	1	Hampshire (FAC)
a22	1	Oxfordshire (BAN)
a23	1	Oxfordshire (BAN)
a24	1	Gloucestershire (SAL)

695

696

S3, Table 3. Modern haplotypes and corresponding GenBank accession numbers from previously published Baker and Hoelzel (2013). * denotes haplotypes shared with the ancient population.

H	n	Location	Corresponding accession numbers from Baker and Hoelzel (2013)
m1*	43	Hampshire, Dorset, Perth, Moray, Durham, N yorkshire, Lanacshire	JX971595, JX971599, JX971604, JX971608, JX971609, JX971615, JX971589
m2*	56	Hampshire, Dorset, Somesret, Berkshire, Moray	JX971590, JX971592
m3	32	Thetford	JX971591
m4*	114	Berkshire, Perth, Durham, N yorkshire, Carlisle, Moray, Ayr, Lancashire	JX971614, JX971612, JX971602, JX971597, JX971593
m5*	12	Perth, Moray	JX971594, JX971596, JX971607
m6*	4	Lancashire	JX971601,JX971598
m7	1	Lancshire	JX971600
m8	1	Perth	JX971603
m9*	9	Perth, Moray, Ayr	JX971605
m10	3	Moray	JX971606
m11	1	N Yorkshire	JX971610
m12	3	Ayr	JX971611, JX971613

S4, Table 4. List of the 388 base pairs length consensus haplotypes obtained for the present analysis after pooling our 23 haplotypes with the 161 haplotypes from Randi et al, (2004), and the 31 haplotypes from Royo et al., (2007). Identification of the sequences collapsing in each haplotype consensus along with the frequencies is detailed. Samples from Randi et al. (2004) and Royo et al., (2007) are identified by their corresponding GenBank accession number, whilst our sequences are identified following S2, Table 2 and S3 Table 3.

Consensus haplotype	Cluster	Frequency observed	Corresponding accession numbers from Randi et al., (2004) and Royo et al., (2007) and/or UK haplotype (as provided in S2 and S3).
C01	Celtic-Iberian	27	AY625732, DQ384673,DQ384672, DQ114774, DQ114772, DQ114771, DQ114769, DQ114768, DQ114767, DQ114765, DQ114747, DQ11474
C02	Celtic-Iberian	16	AY625733, AY625736, DQ384675, DQ384671, DQ384669, DQ114784, DQ114781, DQ114779, DQ114778, DQ114775, DQ114762, DQ114757
C03	Central	18	AY625734, AY625735, DQ384677, DQ384676, DQ114782, DQ114776, DQ114761, DQ114750, DQ114748
C04	West	20	AY625737, AY625807, DQ114773, DQ114770, DQ114766, DQ114752, DQ114751, DQ114745
C05	Central	3	AY625738
C06	East	1	AY625739
C07	East	1	AY625740
C08	East	2	AY625741
C09	East	12	AY625742
C10	Central	7	AY625743
C11	Central	57	AY625744, AY625762, AY625803
C12	Central	77	AY625745, AY625768, AY625816, AY625818, AY625836, AY625853, AY625854

C13	*C.c.italicus*	83	AY625746, AY625766, AY625773,
			AY625775
C14	East	31	AY625747, AY625754, AY625780, AY625784, AY625852, AY625867, DQ384656
C15	East	11	AY625748, AY625883
C16	Central	92	AY625749, AY625772, AY625777, AY625783, AY625827,DQ384688, DQ384686, DQ384684, DQ384683, DQ384664, DQ384663, DQ384657, DQ384655, a3, m2
C17	Central	95	AY625750, DQ384708, DQ384703, DQ384702, DQ384687,DQ384685, DQ384681, DQ384679, DQ114755, DQ114754, a2, m1
C18	Central	4	AY625751, DQ384661, DQ384660, DQ384658
C19	Central	2	AY625753
C20	Central	42	AY625755, m3
C21	West	13	AY625756
C22	East	41	AY625758, AY625842
C23	East	21	AY625759, AY625843, AY625847, AY625877, AY625886
C24	East	24	AY625760, AY625861
C25	East	25	AY625761, DQ384682
C26	Central	1	AY625763
C27	East	1	AY625764
C28	Central	1	AY625765
C29	*C.c.italicus*	1	AY625767
C30	Central	59	AY625769, AY625779, AY625785, AY625796, AY625811, AY625838, AY625848, AY625864, AY625869, a1, m6
C31	Central	1	AY625770
C32	*C.c.italicus*	9	AY625771
C33	*C.c.italicus*	4	AY625774
C34	*C.c.italicus*	8	AY625776, AY625791
C35	Central	1	AY625781

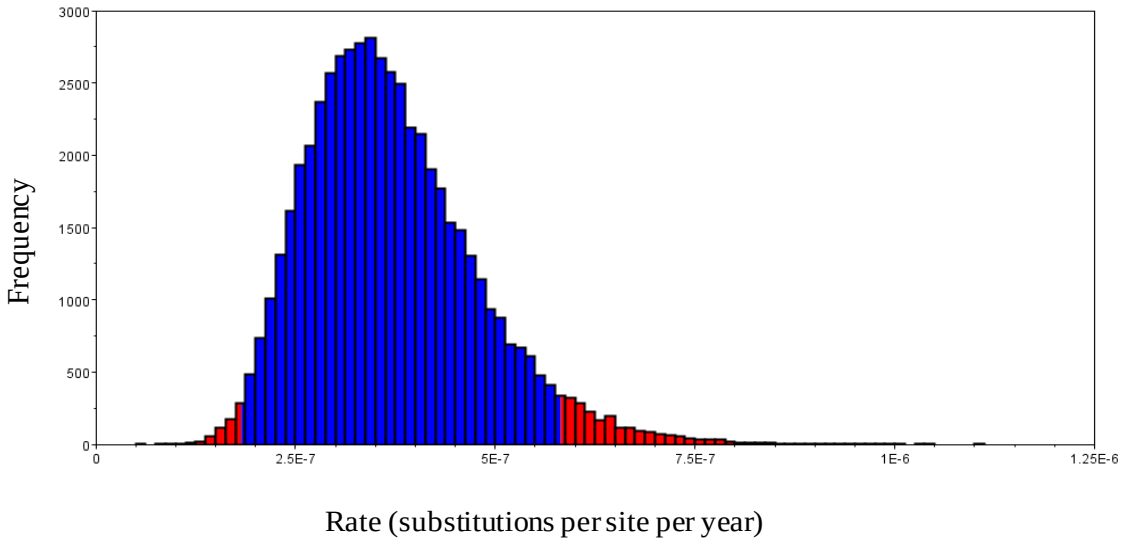
C36	Central	2	AY625786, AY625831
C37	West	1	AY625787
C38	West	2	AY625788
C39	Central	1	AY625789
C40	Central	1	AY625790
C41	Central	1	AY625792
C42	Central	3	AY625793
C43	Central	5	AY625794
C44	Central	1	AY625795
C45	Central	2	AY625798
C46	Central	1	AY625799
C47	West	14	AY625800, AY625801, AY625806
C48	West	2	AY625802
C49	Central	1	AY625804
C50	Celtic-Iberian	13	AY625805,DQ384707, DQ384706, DQ384705, DQ384704, DQ384701, DQ384700, DQ384699, DQ384690, DQ384674, DQ114753, DQ114749
C51	West	1	AY625808
C52	Central	1	AY625809
C53	East	1	AY625810
C54	East	6	AY625812, AY625850, AY625875
C55	East	8	AY625813, AY625859, AY625873
C56	Central	41	AY625814, AY625820, AY625834, AY625837, AY625885, a13, m5
C57	Central	20	AY625815, a25
C58	Central	11	AY625817, AY625857
C59	East	1	AY625819
C60	East	10	AY625821, AY625840, AY625846, AY625870
C61	Central	1	AY625822

C62	West	12	AY625823, AY625824, DQ384668, DQ384667, DQ114783, DQ114777,DQ114764
C63	Central	1	AY625825
C64	Central	2	AY625826
C65	Central	3	AY625828, AY625829
C66	Central	1	AY625830
C67	East	16	AY625833, AY625841, AY625844, AY625851, AY625858, AY625878, AY625880, AY625887
C68	Central	4	AY625835, AY625856
C69	East	1	AY625845
C70	Central	2	AY625849
C71	Central	6	AY625855, AY625865, AY625872, AY625876
C72	East	1	AY625860
C73	East	3	AY625862, AY625881
C74	East	4	AY625863, AY625871, AY625884
C75	Central	11	AY625866, a17, m9
C76	Central	1	AY625874
C77	East	1	AY625882
C78	Central	1	AY625888
C79	Central	9	AY625890, DQ384697, DQ384696, DQ384694, DQ384691, DQ384689
C80	West	7	AY625891, DQ384652, DQ384651, DQ384649, DQ384648, DQ384646, DQ384641
C81	West	1	AY625892
C82	Central	2	a 4
C83	Central	2	a 6
C84	Central	4	a 7
C85	Central	3	a 8
C86	Central	1	a 9
C87	Central	2	a 10

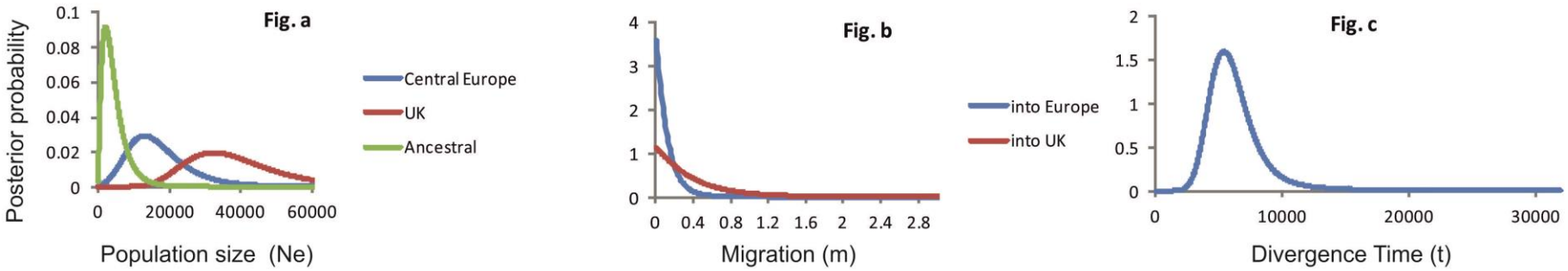
C88	Central	3	a 11
C89	Central	1	a 12
C90	Central	1	a 14
C91	Central	1	a 15
C92	Central	125	a16, m4
C93	Central	1	a 18
C94	Central	1	a 19
C95	Central	1	a 20
C96	Central	1	a 21
C97	Central	1	a 22
C98	Central	1	a 23
C99	Central	1	a 24
C100	Central	1	m 7
C101	Central	1	m8
C102	Central	3	m10
C103	Central	1	m11
C104	Central	3	m12
C105	West	12	DQ384698, DQ384693, DQ384692, DQ384666, DQ384665, DQ384654, DQ384653, DQ384647, DQ384645, DQ384644, DQ384643, DQ384642
C106	Central	1	DQ384695
C107	Central	1	DQ384680
C108	Celtic-Iberian	1	DQ384678
C109	Celtic-Iberian	1	DQ384670
C110	West	2	DQ384662,DQ384659
C111	West	1	DQ384650
C112	West	1	DQ384640
C113	Celtic-Iberian	1	DQ114780

C114	West	1	DQ114763
C115	Central	4	DQ114760, DQ114759, DQ114758, DQ114756

706
707



S1, Figure 1. BEAST output for the roe deer substitution rate (3.69 x 10 -7) estimated under a BSP population size model from 86 roe deer stratigraphic date samples from the UK.



714

715 S2, Figure 2. The posterior probabilities of demographic model parameter estimates based on an estimate of μ (see text) where: Ne = effective
716 population size of ancient UK, central Europe and the ancestral population (Fig a); m = average number of migrants per 1000 generations per
717 gene copy into ancient UK and central European populations (Fig b) and t = divergence time in years between ancient UK and central Europe
718 (Fig c).

719

720